

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 1 104 155 A2

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
30.05.2001 Bulletin 2001/22

(51) Int Cl.7: H04M 3/42, H04M 3/493,
G10L 15/00, H04M 7/00

(21) Application number: 00310429.6

(22) Date of filing: 24.11.2000

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: 24.11.1999 US 167290 P
21.08.2000 US 643558

(71) Applicant: Phone.Com Inc.
Redwood City, CA 94063 (US)

(72) Inventor: King, Peter F.
California 94019 (US)

(74) Representative: Ablett, Graham Keith et al
Ablett & Stebbing,
Caparo House,
101-103 Baker Street
London W1M 1FD (GB)

(54) Voice recognition based user interface for wireless devices

(57) The present invention relates to a wireless communication system (100) that utilizes a remote voice recognition server system (109) to translate voice input received from serviced mobile devices (102, 103) into a symbolic data file (e.g. alpha-numeric or control characters) that can be processed by the mobile devices. Initially, a voice communication channel (126) is established between the serviced mobile device and the voice recognition server. A user of the mobile device then be-

gins speaking in a fashion that may be detected by the voice recognition server system. Upon detecting the user's speech, the voice recognition server system uses a voice recognition application to translate the speech into a symbolic data file, which is then forwarded to the user through a separate data communication channel (128, 130). The user, upon receiving the symbolic data file at the mobile device, reviews and edits the content and further utilizes the file as desired.

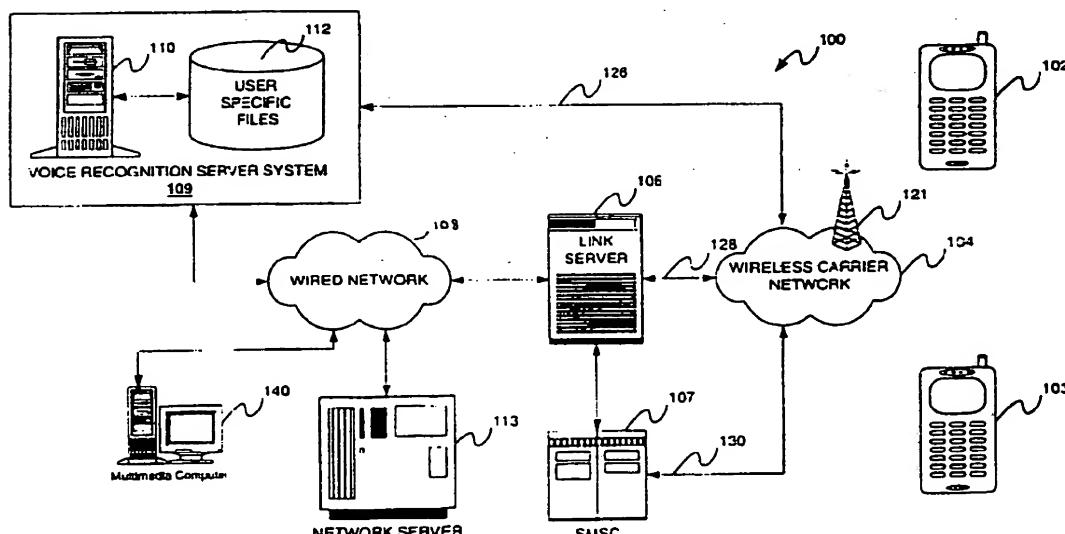


Figure 1

Description**BACKGROUND OF THE INVENTION****Field of Invention**

[0001] This invention generally relates to data communications, and in particular to a two-way wireless communication device that utilizes network based voice recognition resources to augment the local user interface.

Discussion of Related Art

[0002] The use of hypertext based technologies has spread to the domain of wireless communication systems. Two-way wireless communication devices, also described as mobile devices herein, and wireless network protocols have been designed to permit interactive access to remote information services (e.g. commercial databases, email, on-line shopping), through a variety of wireless and wire-line networks, most notably the Internet and private networks.

[0003] Many mobile devices (e.g. cellular telephones) are mass-market consumer oriented-devices. Their user interface should thus be simple and easy to use without limiting the functionality of the device. Currently, the primary method of data entry for most mobile devices is a keypad that is relatively inefficient when used to input lengthy alphanumeric character strings. Due to size constraints and cost considerations, the keypads of these mobile devices are not a particularly user friendly interface for drafting messages requiring substantial user input (e.g. email messages). Keypads of this type usually have between 12 and 24 keys, a sufficient number for numeric inputs but very inefficient when dealing with the alphanumeric data entries required for network capable devices.

[0004] A user requesting information from the Internet generally navigates the World Wide Web using a browser. For example, a user requesting information on *Stanford University* using Infoseek™ as the search engine would have to input the following string:

"<http://www.infoseek.com>" followed by "Stanford University"

[0005] The search string listed above includes over 40 characters. A user would have no problem inputting a string of this type using a standard desktop computer keyboard and browser (e.g. Netscape or Explorer). However, the same user operating the keypad of a mobile device to input the same string would be severely hampered by the compact keypad and the close spacing between the keys.

[0006] One of the common uses of the Internet is email. A user who desires to send an email message having the size of the paragraph above would have to input over 400 characters. Using the standard keyboard of a desktop computer, a user may be able to input that

number of characters in less than two minutes (assuming the user could type with an average degree of skill). Inputting the same number of keystrokes on the keypad of a mobile device could take considerably longer and become very tedious and prone to error.

[0007] Recent advances in voice recognition (VR) technology and increases in hardware capabilities are making the development of voice-recognition-based user interfaces for desktop systems commercially viable. VR technology takes spoken words and translates them into a format, which can easily be manipulated and displayed by digital systems. There have been efforts to equip compact mobile devices with VR technology, however, these efforts have generally required costly device modifications such as extra components (e.g. a DSP chip) or increased processing and storage capability. A typical cellular phone has computational resources equivalent to less than one percent of what is provided in a typical desktop or portable computer. A phone of this type running a scaled down VR application would only be able to recognize a small-predefined group of spoken words without modifying the device components.

[0008] Voice recognition software currently available for desktop and laptop computers (e.g. Naturally Speaking from Dragon System, Inc., PlainTalk™ from Apple Computer, ViaVoice 98™ from IBM and FreeSpeech 98™ from Philips Talk) generally costs between \$39.00 and several hundred dollars per license. This would represent a significant portion of the costs of a mobile device equipped with a comparable software application.

[0009] Placing a voice recognition software application in each mobile device and modifying its hardware components to run that application creates a financial disincentive for the handset manufacturers to incorporate VR features in their devices. These modifications would add considerable cost to the final price of the mobile device, possibly pricing them out of the target price range (e.g. \$150.00) usually occupied by mass-market mobile devices (e.g. cellular telephones).

[0010] In terms of hardware resources, these applications can require up to 60 Mbytes of memory for each language supported. Additionally most of the commercially available voice recognition software applications are designed to function on systems having relatively fast processors (e.g. 133 MHZ Pentium processor).

[0011] There is thus a great need for apparatuses and methods that enable mobile devices to interact in a more efficient manner with digital computer networks. The ability to utilize voice recognition services in conjunction with the standard mobile device user interface (e.g. a phone keypad), without having to significantly modify hardware resources or costs, would dramatically improve the usability and commercial viability of network capable mobile devices having limited resources.

SUMMARY OF THE INVENTION

[0012] The present invention relates to a wireless communication system that utilizes a remote voice recognition server system to translate voice input received from mobile devices into a symbolic data file (e.g. alphanumeric or control characters) that can be processed by the mobile devices. The translation process begins by establishing a voice communication channel between a mobile device and the voice recognition server. A user of the mobile device then begins speaking in a fashion that may be detected by the voice recognition server system. Upon detecting the user's speech, the voice recognition server system translates the speech into a symbolic data file, which is then forwarded to the user through a separate data communication channel. The user, upon receiving the symbolic data file at the mobile device, reviews and edits the content of the symbolic data file and further utilizes the file as desired. For example a user could use the symbolic data file to fill in fields in an email or a browser request field.

[0013] The invention can be implemented in numerous ways, including as a method, an apparatus or device, a user interface, a computer readable memory and a system. Several embodiments of the invention are discussed below.

[0014] According to one embodiment, the present invention is a method for obtaining voice recognition services for a mobile device not having the resources and/or software for performing voice recognition processing locally. The method comprises using local applications resident within the mobile device to establish and coordinate a voice channel between the subject mobile device and a remote server system running a voice recognition application (referred to herein as a voice recognition server system).

[0015] Upon establishment of the voice channel the user of the subject mobile device is queued to begin speaking into the microphone of the mobile device (e.g. a cellular phone). Voiced input received at the voice recognition server system, as a result of this interaction, is converted into a symbolic data file. This process may be assisted by previously stored user specific data files. The symbolic data file is then forwarded back to the originating mobile device or a designated third party device through a separately established and coordinated data communication channel. The symbolic data file may be used to interact with local applications on the mobile device or to interact with network resources (e.g. servers on the Internet or a private network).

[0016] Other objects and advantages, together with the foregoing are attained in the exercise of the invention in the following description and accompanying drawings.

BRIEF DESCRIPTION OF DRAWINGS

[0017] The present invention will be readily under-

stood by the following detailed description in conjunction with the accompanying drawings, wherein the reference numerals illustrate the structural elements, and in which:

5 Figure 1 illustrates a schematic configuration in which the present invention may be practised; Figure 2A depicts display and user interface components of a typical voice capable mobile device; Figure 2B illustrates a functional block diagram of an exemplary voice capable mobile device; Figure 3 illustrates a functional block diagram of a link server device according to a preferred embodiment of the present invention; Figure 4 is a schematic diagram showing exemplary processing stages for the voice recognition server in accordance with an exemplary embodiment of the present invention; Figure 5 shows representative screen displays, which illustrate operations relating to the interaction of a mobile device with a voice recognition server system; Figure 6 illustrates a process flowchart from the perspective of the mobile device according to an embodiment the present invention; and Figure 7 illustrates a process flowchart from the perspective of the voice recognition server according to an embodiment the present invention.

DETAILED DESCRIPTION OF THE INVENTION

30 [0018] In the following detailed description of the present invention, numerous specific details are set forth in order to provide a thorough understanding of the present invention. However, it will become obvious to those skilled in the art that the present invention may be practised without these specific details. In other instances, well known methods, procedures, components, and circuitry have not been described in detail to avoid unnecessarily obscuring aspects of the present invention.

35 [0019] The detailed description of the present invention in the following are presented largely in terms of procedures, steps, logic blocks, processing, and other symbolic representations that resemble data processing devices coupled to networks. These process descriptions and representations are the means used by those experienced or skilled in the art to most effectively convey the substance of their work to others skilled in the art.

40 [0020] The invention pertains to systems and methods, which enable a mobile device to access voice recognition services from a networked voice recognition server system. According to one embodiment of the present invention, voice recognition services are accessed by establishing a voice channel between the user of a mobile device desiring voice recognition services and a networked voice recognition server system.

45 [0021] Once a voice channel is established, the user of the mobile device is given a queue to begin speaking when the voice recognition server system is ready to re-

ceive a speech signal. The received speech signal is processed by the voice recognition server system using voice recognition techniques well known in the art (e.g. template matching, Fourier transforms or linear predictive coding (LPC)) and a symbolic data file is generated. [0021] A symbolic data file is a file containing a plurality of letters, phonemes, words, figures, objects, functions, control characters or other conventional marks designating an object, quantity, operation, function, phoneme, word, phrase or any combination thereof having some relationship to the received speech signal as interpreted by the voice recognition system. Voice recognition systems generally use voice templates, Fourier Transform coding, or a linear predictive coding scheme to map the voiced input components to pre-stored symbolic building blocks. Examples of symbolic data files include ASCII files and binary data files.

[0022] To facilitate a description of the present invention, it is useful to recite some of the features of a communication system in which the invention may be practised. Figures 1 through 4 provide an overview of the principal system components.

[0023] Referring to Figure 1 a block diagram of a typical communication system according to one embodiment of the present invention is displayed. Mobile devices 102 and 103 receive phone calls through a voice communication channel and hypermedia information (e.g. Hyper Text Markup Language (HTML) documents, Compact Hypertext Transport Protocol (cHTML) documents, Extensible Markup Language (XML) documents, Handheld Device Markup Language (HDML) documents, or Wireless Markup Language (WML) documents, or other similar data types) from remote server devices through broad-band and narrow-band (e.g. SMS) data communication channels which may include link server device 106 and Short Message Service center (SMSC) 107.

[0024] Mobile devices 102 and 103 each have a display and a user interface. Additionally, mobile devices 102 and 103 may have a micro-browsers (e.g. a micro-browser from Phone.com, Inc. 800 Chesapeake Drive, Redwood City, CA, 94063) stored in a local memory (also referred to as a client module) which enables the device to process hypermedia information received from remote server devices.

[0025] As shown in Figure 1, mobile devices 102 and 103 may be coupled to link server device 106 through a wireless carrier network 104 (also referred to herein as a wireless network). Mobile devices 102 and 103 may be taken from a group, which includes mobile phones, palm sized computing devices and personal digital assistants with voice transmission and/or reception capabilities. Voice capabilities are defined as the capabilities equipped in a mobile device that allow a user to communicate voice based information to and from remote destinations (e.g. to another user or a device).

[0026] Access to the voice communication channel generally requires that the user and/or device be recog-

nized by wireless carrier network 104. Network recognition involves the exchange of identification information between a subject mobile device and wireless carrier network 104. Generally, the identification information for

5 the user and/or mobile device in question is stored in the memory of the device and is transmitted automatically when the user attempts to access the network.

[0027] Wireless carrier network 104 may be any of the 10 well known wireless communication networks (e.g. cellular digital packet data (CDPD) network, Global System for Mobile Communication (GSM) network, Code Division Multiple Access (CDMA) network, Personal Handy Phone System (PHS) or Time Division Multiple Access (TDMA) network). Link server device 106 is further coupled to a wired network 108 to which a voice recognition server system 109 and a plurality of networked servers 15 represented by network server 113 are coupled.

[0028] Voice recognition server system 109 is comprised of a server device 110 and storage facilities 112 20 capable of storing, among other things, user specific files associated with a plurality of user's serviced by a carrier entity. The user specific files are utilized in conjunction with voice recognition processing and in one embodiment are part of the present invention.

25 [0029] Examples of user specific files might include user specific speech templates, one or more user specified language dictionaries (e.g. French, English, German or Cantonese) and one or more user specific dictionaries or lists of an individual user's frequently used words. These files may be uploaded and managed using a networked multimedia computer (e.g. multimedia computer 140) or through the user interface of the serviced mobile device. For example, voice templates are 30 generated by having the user read a pre-determined script into a voice-enabled device. User preferences (e.g. languages of choice) may be input using menu selection screens presented to the user on the display of the mobile device or another device connected to the voice recognition server system via a wired network.

35 [0030] For simplicity, antenna 121 represents a wireless carrier infrastructure that generally comprises a base station and an operations and maintenance center. The base station controls radio or telecommunication links with mobile devices 102 and 103. The operations 40 and maintenance center comprises a mobile switching center, which switches calls between the mobile devices and other fixed or mobile network users. Further the operations and maintenance center manages mobile account services, such as authentication, and oversees 45 the proper operation and setup of the wireless network. Each of the hardware components and processes in carrier infrastructure 121 are known to those skilled in the art and will not be described herein to avoid unnecessarily obscuring aspects of the present invention.

50 [0031] The communication protocols used by airtel 104 may, for example, be Wireless Access Protocol (WAP) or Handheld Device Transport Protocol (HDTP). Wired network 108 is a land-based network that may be

the Internet, a private network or a data network of any private network. Typically the communication protocol supporting landnet 118 may be Transmission Control Protocol (TCP/IP), Hypertext Transport Protocol (HT-TP), or Secure Hypertext Transport Protocol (sHTTP). [0032] Link server device 106 and network server 113 are typically computer work stations, for example, a SPARC station from Sun Microsystems Inc. (<http://www.sun.com>) with networking libraries and Internet connectivity. Network server 113 is representative of a plurality of networked servers coupled to landnet 108 and is capable of providing access to hypermedia information including information for mobile devices 102 and 103.

[0033] Link server device 106 is depicted as a stand alone device and therefore is often referred to as a network gateway or wireless data server. Link server 106 can be configured to operate as a bridge between wireless network 104 and wired network 108. It should be pointed out that the functions of link server device 106 may be performed by other server devices connected to wired network 108 with hardware well known in the art providing the connection between wireless network 104 and wired network 108.

[0034] The voice communication channel previously described is generally represented by voice channel 126. This communication channel is generally established and coordinated using the infrastructure and procedures generally known in the art for setting up a phone call.

[0035] There are generally two types of data communication channels providing service to mobile devices 102 and 103. Data communication channel 128 is representative of a wideband data communication channel. Data communication channel 130 is representative of a narrowband data communication channel e.g. (a Short Message Communication (SMS) service channel). Either of these data communication paths can be used to deliver data to and from mobile devices 102 and 103.

[0036] According to the preferred embodiment of the present invention a mobile device (e.g. mobile device 102 or 103) desiring to receive voice recognition services from voice recognition server system 109, first establishes a voice channel generally represented by voice channel 126. The contact information for voice recognition server system 109 (e.g. a phone number or a uniform resource indicator (URI)) may be embedded in software loaded on the mobile device, retrieved from link server device 106 or input by the user directly.

[0037] Once a voice channel is established between the requesting mobile device and voice recognition server system 109, user information is forwarded to the voice recognition server system. This allows previously stored user specific files for the requesting mobile device to be accessed and utilized. The user information may be transmitted on a separate data communication channel (e.g. data communication channels 128 or 130) or input by the user. The user specific files generally provide for features specific to a particular user account.

For example, the user can specify one or more languages of choice for voice recognition processing.

[0038] Once the user specific files for the subject mobile device/user are retrieved, the user is prompted to provide a voiced input (e.g. begin speaking). It is important to note at this point that the user may utilize the user interface of the mobile device (e.g. a phone keypad) while utilizing voice recognition services. When the user has completed their input interaction (voice and physical input) with the mobile device, an indication may be provided by the user (voiced or key input) to conclude the input session. Voice recognition server system 109 then converts the voiced input into a symbolic data file, which can be forwarded to the requesting mobile device via link server 106.

[0039] As previously stated, the symbolic data file is a file containing a plurality of letters, phonemes, words, figures, objects, functions, control characters or other conventional marks designating an object, quantity, operation, function, phoneme, word, phrase or any combination thereof having some relationship to the received speech signal as interpreted by the voice recognition system. Voice recognition systems generally use voice templates, Fourier Transform coding, or a linear predictive coding scheme to map the voiced input components to pre-stored symbolic building blocks. Examples of symbolic data files include ASCII files and binary data files.

[0040] The symbolic data file may initially be forwarded to link server device 106, which may perform additional processing prior to forwarding the symbolic data file to the requesting mobile device via wideband channel 128 or narrowband channel 130. The user of the mobile device may then review the received symbolic data file and utilize it as desired.

[0041] The accuracy of the voice recognition application used by voice recognition server system 109 will vary depending on the translation methodology used and the size and language of the language dictionaries used. Generally, speaker dependent methodologies (e.g. template matching) have accuracy's as high as 98 percent and speaker-independent methodologies (e.g. Fourier transforms and linear predictive coding (LPC)) have accuracy's in the range of 90 to 95 percent (www.hitl.washington.edu - *Voice Recognition*, Jim Baumann).

[0042] In accordance with the principles of the present invention, users of mobile devices (e.g. mobile devices 102 and 103) may access voice recognition services on those mobile devices without the significant hardware or software modifications that might be required if the voice recognition application were executed by the device. Additionally, since the software performing voice recognition processing is resident on an accessible remote server device with superior processing speed (as compared to that of the mobile device) and large storage capacity, the user of the device can be provided with the functionality and resources associated with a full featured voice recognition application. For example, the

voice recognition application may have access to large language dictionaries, selectable language dictionaries for multiple languages and user specific files (e.g. voice templates and user customized dictionaries and lists).

[0043] Figure 2A depicts an exemplary mobile device 200 that may correspond to one of mobile devices (102 or 103) in Figure 1. Mobile device 200 includes a display screen 204, an extended phone-styled keypad 210, cursor navigation keys 222 and 224, a pair of softkeys 208A and 208B, an earpiece 212A and a microphone 212B. Screen display 204 is typically a Liquid Crystal Display (LCD) screen capable of displaying textual information and certain graphics. Extended phone keypad 210 includes, preferably, a regular phone keypad with additional keys providing additional characters (e.g. a space) and functions (e.g. back or clear).

[0044] Cursor navigation keys 222 and 224 allow a user to reposition a cursor or an element indicator 216, for example, to activate one of the applications displayed on screen display 204. Generic keys 208A and 208B are typically used to perform application specific functions as indicated by softkey function identifiers 214 and 215. It should be understood, by those having ordinary skill in the art, that having a regular phone keypad is not a requirement to practice the present invention. Some mobile devices sometimes have no physical keys at all, such as those palm-sized computing devices that use soft keys or icons as an input mechanism.

[0045] Upon establishing a communication session with an associated link server device (e.g., link server device 106 of Figure 1) mobile device 200 typically receives one or more markup language card decks to assist the user with device interactions. Depending on the implementation preference, the markup language card decks, alternatively referred to as screen descriptive commands files, may be in a markup language that includes, but is not limited to, Handheld Device Markup Language (HDML), Hypertext Markup Language (HTML), compact HTML, Wireless Markup Language (WML), Standard Generalized Markup Language (SGML) or Extensible Markup Language (XML). Alternatively, the data file may be a stripped, compressed, compiled or converted version of a corresponding markup file.

[0046] The text appearing on LCD screen 204 in Figure 2A is an example of such a display screen. In this example the user is offered a choice of the following selections:

- 1) Bookmarks
- 2) Search Int.
- 3) Email
- 4) News

Each of the selections is typically linked to a resource on the network or to a local software application. A user may make a selection from the menu above using navigation keys 222 and 224 with the user's selection indi-

cated by element indicator 216. This same method may be utilized to provide user prompts for interacting with remote server devices (e.g. voice recognition server system 109 of Figure 1).

[0047] Referring now to Figure 2B, a more detailed description of mobile device 250, which may be mobile devices 102 or 103 of Figure 1 and 200 of Figure 2A, is provided. Mobile device 250 includes a Wireless Control Protocol (WCP) interface 252 that couples to a carrier

10 wireless network 104 to receive incoming and outgoing signals. Device identifier (ID) storage 254 stores and supplies a device ID to WCP interface 252 for the purpose of identifying mobile device 250 to outside entities (e.g. link server device 106 of Figure 1). The device ID

15 is a specific code that is associated with mobile device 250 and directly corresponds to the device ID in an associated user account typically provided in an associated link server device (e.g. 106 of Figure 1).

[0048] Mobile device 250 includes a processor 268, encoder/decoder circuitry 264, working memory 258 and a client module 256. Client module 256 is representative of software components loaded on or into device memory resources, which performs many of the processing tasks performed by mobile device 250 including; establishing a communication session with a link server device via wireless carrier network 104, operating and maintaining local applications, displaying information on a display screen 260 of mobile device 250, and receiving user input from keypad 262. Client module

25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995 1000 1005 1010 1015 1020 1025 1030 1035 1040 1045 1050 1055 1060 1065 1070 1075 1080 1085 1090 1095 1100 1105 1110 1115 1120 1125 1130 1135 1140 1145 1150 1155 1160 1165 1170 1175 1180 1185 1190 1195 1200 1205 1210 1215 1220 1225 1230 1235 1240 1245 1250 1255 1260 1265 1270 1275 1280 1285 1290 1295 1300 1305 1310 1315 1320 1325 1330 1335 1340 1345 1350 1355 1360 1365 1370 1375 1380 1385 1390 1395 1400 1405 1410 1415 1420 1425 1430 1435 1440 1445 1450 1455 1460 1465 1470 1475 1480 1485 1490 1495 1500 1505 1510 1515 1520 1525 1530 1535 1540 1545 1550 1555 1560 1565 1570 1575 1580 1585 1590 1595 1600 1605 1610 1615 1620 1625 1630 1635 1640 1645 1650 1655 1660 1665 1670 1675 1680 1685 1690 1695 1700 1705 1710 1715 1720 1725 1730 1735 1740 1745 1750 1755 1760 1765 1770 1775 1780 1785 1790 1795 1800 1805 1810 1815 1820 1825 1830 1835 1840 1845 1850 1855 1860 1865 1870 1875 1880 1885 1890 1895 1900 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000 2005 2010 2015 2020 2025 2030 2035 2040 2045 2050 2055 2060 2065 2070 2075 2080 2085 2090 2095 2100 2105 2110 2115 2120 2125 2130 2135 2140 2145 2150 2155 2160 2165 2170 2175 2180 2185 2190 2195 2200 2205 2210 2215 2220 2225 2230 2235 2240 2245 2250 2255 2260 2265 2270 2275 2280 2285 2290 2295 2300 2305 2310 2315 2320 2325 2330 2335 2340 2345 2350 2355 2360 2365 2370 2375 2380 2385 2390 2395 2400 2405 2410 2415 2420 2425 2430 2435 2440 2445 2450 2455 2460 2465 2470 2475 2480 2485 2490 2495 2500 2505 2510 2515 2520 2525 2530 2535 2540 2545 2550 2555 2560 2565 2570 2575 2580 2585 2590 2595 2600 2605 2610 2615 2620 2625 2630 2635 2640 2645 2650 2655 2660 2665 2670 2675 2680 2685 2690 2695 2700 2705 2710 2715 2720 2725 2730 2735 2740 2745 2750 2755 2760 2765 2770 2775 2780 2785 2790 2795 2800 2805 2810 2815 2820 2825 2830 2835 2840 2845 2850 2855 2860 2865 2870 2875 2880 2885 2890 2895 2900 2905 2910 2915 2920 2925 2930 2935 2940 2945 2950 2955 2960 2965 2970 2975 2980 2985 2990 2995 3000 3005 3010 3015 3020 3025 3030 3035 3040 3045 3050 3055 3060 3065 3070 3075 3080 3085 3090 3095 3100 3105 3110 3115 3120 3125 3130 3135 3140 3145 3150 3155 3160 3165 3170 3175 3180 3185 3190 3195 3200 3205 3210 3215 3220 3225 3230 3235 3240 3245 3250 3255 3260 3265 3270 3275 3280 3285 3290 3295 3300 3305 3310 3315 3320 3325 3330 3335 3340 3345 3350 3355 3360 3365 3370 3375 3380 3385 3390 3395 3400 3405 3410 3415 3420 3425 3430 3435 3440 3445 3450 3455 3460 3465 3470 3475 3480 3485 3490 3495 3500 3505 3510 3515 3520 3525 3530 3535 3540 3545 3550 3555 3560 3565 3570 3575 3580 3585 3590 3595 3600 3605 3610 3615 3620 3625 3630 3635 3640 3645 3650 3655 3660 3665 3670 3675 3680 3685 3690 3695 3700 3705 3710 3715 3720 3725 3730 3735 3740 3745 3750 3755 3760 3765 3770 3775 3780 3785 3790 3795 3800 3805 3810 3815 3820 3825 3830 3835 3840 3845 3850 3855 3860 3865 3870 3875 3880 3885 3890 3895 3900 3905 3910 3915 3920 3925 3930 3935 3940 3945 3950 3955 3960 3965 3970 3975 3980 3985 3990 3995 4000 4005 4010 4015 4020 4025 4030 4035 4040 4045 4050 4055 4060 4065 4070 4075 4080 4085 4090 4095 4100 4105 4110 4115 4120 4125 4130 4135 4140 4145 4150 4155 4160 4165 4170 4175 4180 4185 4190 4195 4200 4205 4210 4215 4220 4225 4230 4235 4240 4245 4250 4255 4260 4265 4270 4275 4280 4285 4290 4295 4300 4305 4310 4315 4320 4325 4330 4335 4340 4345 4350 4355 4360 4365 4370 4375 4380 4385 4390 4395 4400 4405 4410 4415 4420 4425 4430 4435 4440 4445 4450 4455 4460 4465 4470 4475 4480 4485 4490 4495 4500 4505 4510 4515 4520 4525 4530 4535 4540 4545 4550 4555 4560 4565 4570 4575 4580 4585 4590 4595 4600 4605 4610 4615 4620 4625 4630 4635 4640 4645 4650 4655 4660 4665 4670 4675 4680 4685 4690 4695 4700 4705 4710 4715 4720 4725 4730 4735 4740 4745 4750 4755 4760 4765 4770 4775 4780 4785 4790 4795 4800 4805 4810 4815 4820 4825 4830 4835 4840 4845 4850 4855 4860 4865 4870 4875 4880 4885 4890 4895 4900 4905 4910 4915 4920 4925 4930 4935 4940 4945 4950 4955 4960 4965 4970 4975 4980 4985 4990 4995 5000 5005 5010 5015 5020 5025 5030 5035 5040 5045 5050 5055 5060 5065 5070 5075 5080 5085 5090 5095 5100 5105 5110 5115 5120 5125 5130 5135 5140 5145 5150 5155 5160 5165 5170 5175 5180 5185 5190 5195 5200 5205 5210 5215 5220 5225 5230 5235 5240 5245 5250 5255 5260 5265 5270 5275 5280 5285 5290 5295 5300 5305 5310 5315 5320 5325 5330 5335 5340 5345 5350 5355 5360 5365 5370 5375 5380 5385 5390 5395 5400 5405 5410 5415 5420 5425 5430 5435 5440 5445 5450 5455 5460 5465 5470 5475 5480 5485 5490 5495 5500 5505 5510 5515 5520 5525 5530 5535 5540 5545 5550 5555 5560 5565 5570 5575 5580 5585 5590 5595 5600 5605 5610 5615 5620 5625 5630 5635 5640 5645 5650 5655 5660 5665 5670 5675 5680 5685 5690 5695 5700 5705 5710 5715 5720 5725 5730 5735 5740 5745 5750 5755 5760 5765 5770 5775 5780 5785 5790 5795 5800 5805 5810 5815 5820 5825 5830 5835 5840 5845 5850 5855 5860 5865 5870 5875 5880 5885 5890 5895 5900 5905 5910 5915 5920 5925 5930 5935 5940 5945 5950 5955 5960 5965 5970 5975 5980 5985 5990 5995 6000 6005 6010 6015 6020 6025 6030 6035 6040 6045 6050 6055 6060 6065 6070 6075 6080 6085 6090 6095 6100 6105 6110 6115 6120 6125 6130 6135 6140 6145 6150 6155 6160 6165 6170 6175 6180 6185 6190 6195 6200 6205 6210 6215 6220 6225 6230 6235 6240 6245 6250 6255 6260 6265 6270 6275 6280 6285 6290 6295 6300 6305 6310 6315 6320 6325 6330 6335 6340 6345 6350 6355 6360 6365 6370 6375 6380 6385 6390 6395 6400 6405 6410 6415 6420 6425 6430 6435 6440 6445 6450 6455 6460 6465 6470 6475 6480 6485 6490 6495 6500 6505 6510 6515 6520 6525 6530 6535 6540 6545 6550 6555 6560 6565 6570 6575 6580 6585 6590 6595 6600 6605 6610 6615 6620 6625 6630 6635 6640 6645 6650 6655 6660 6665 6670 6675 6680 6685 6690 6695 6700 6705 6710 6715 6720 6725 6730 6735 6740 6745 6750 6755 6760 6765 6770 6775 6780 6785 6790 6795 6800 6805 6810 6815 6820 6825 6830 6835 6840 6845 6850 6855 6860 6865 6870 6875 6880 6885 6890 6895 6900 6905 6910 6915 6920 6925 6930 6935 6940 6945 6950 6955 6960 6965 6970 6975 6980 6985 6990 6995 7000 7005 7010 7015 7020 7025 7030 7035 7040 7045 7050 7055 7060 7065 7070 7075 7080 7085 7090 7095 7100 7105 7110 7115 7120 7125 7130 7135 7140 7145 7150 7155 7160 7165 7170 7175 7180 7185 7190 7195 7200 7205 7210 7215 7220 7225 7230 7235 7240 7245 7250 7255 7260 7265 7270 7275 7280 7285 7290 7295 7300 7305 7310 7315 7320 7325 7330 7335 7340 7345 7350 7355 7360 7365 7370 7375 7380 7385 7390 7395 7400 7405 7410 7415 7420 7425 7430 7435 7440 7445 7450 7455 7460 7465 7470 7475 7480 7485 7490 7495 7500 7505 7510 7515 7520 7525 7530 7535 7540 7545 7550 7555 7560 7565 7570 7575 7580 7585 7590 7595 7600 7605 7610 7615 7620 7625 7630 7635 7640 7645 7650 7655 7660 7665 7670 7675 7680 7685 7690 7695 7700 7705 7710 7715 7720 7725 7730 7735 7740 7745 7750 7755 7760 7765 7770 7775 7780 7785 7790 7795 7800 7805 7810 7815 7820 7825 7830 7835 7840 7845 7850 7855 7860 7865 7870 7875 7880 7885 7890 7895 7900 7905 7910 7915 7920 7925 7930 7935 7940 7945 7950 7955 7960 7965 7970 7975 7980 7985 7990 7995 8000 8005 8010 8015 8020 8025 8030 8035 8040 8045 8050 8055 8060 8065 8070 8075 8080 8085 8090 8095 8100 8105 8110 8115 8120 8125 8130 8135 8140 8145 8150 8155 8160 8165 8170 8175 8180 8185 8190 8195 8200 8205 8210 8215 8220 8225 8230 8235 8240 8245 8250 8255 8260 8265 8270 8275 8280 8285 8290 8295 8300 8305 8310 8315 8320 8325 8330 8335 8340 8345 8350 8355 8360 8365 8370 8375 8380 8385 8390 8395 8400 8405 8410 8415 8420 8425 8430 8435 8440 8445 8450 8455 8460 8465 8470 8475 8480 8485 8490 8495 8500 8505 8510 8515 8520 8525 8530 8535 8540 8545 8550 8555 8560 8565 8570 8575 8580 8585 8590 8595 8600 8605 8610 8615 8620 8625 8630 8635 8640 8645 8650 8655 8660 8665 8670 8675 8680 8685 8690 8695 8700 8705 8710 8715 8720 8725 8730 8735 8740 8745 8750 8755 8760 8765 8770 8775 8780 8785 8790 8795 8800 8805 8810 8815 8820 8825 8830 8835 8840 8845 8850 8855 8860 8865 8870 8875 8880 8885 8890 8895 8900 8905 8910 8915 8920 8925 8930 8935 8940 8945 8950 8955 8960 8965 8970 8975 8980 8985 8990 8995 9000 9005 9010 9015 9020 9025 9030 9035 9040 9045 9050 9055 9060 9065 9070 9075 9080 9085 9090 9095 9100 9105 9110 9115 9120 9125 9130 9135 9140 9145 9150 9155 9160 9165 9170 9175 9180 9185 9190 9195 9200 9205 9210 9215 9220 9225 9230 9235 9240 9245 9250 9255 9260 9265 9270 9275 9280 9285 9290 9295 9300 9305 9310 9315 9320 9325 9330 9335 9340 9345 9350 9355 9360 9365 9370 9375 9380 9385 9390 9395 9400 9405 9410 9415 9420 9425 9430 9435 9440 9445 9450 9455 9460 9465 9470 9475 9480 9485 9490 9495 9500 9505 9510 9515 9520 9525 9530 9535 9540 9545 9550 9555 9560 9565 9570 9575 9580 9585 9590 9595 9600 9605 9610 9615 9620 9625 9630 9635 9640 9645 9650 9655 9660 9665 9670 9675 9680 9685 9690 9695 9700 9705 9710 9715 9720 9725 9730 9735 9740 9745 9750 9755 9760 9765 9770 9775 9780 9785 9790 9795 9800 9805 9810 9815 9820 9825 9830 9835 9840 9845 9850 9855 9860 9865 9870 9875 9880 9885 9890 9895 9900 9905 9910 9915 9920 9925 9930 9935 9940 9945 9950 9955 9960 9965 9970 9975 9980 9985 9990 9995 9999

device running the voice recognition application, if not already stored, and the process would proceed as described above. This example is provided for purposes of illustration and should not be interpreted as limiting the scope of the present invention.

[0052] Figure 3 schematically illustrates the principle components of link server device 340, which may correspond to link server device 106 of Figure 1. Link server device 340 is a server computer that operates as a network gateway between wired network 300 and wireless network 320. To avoid obscuring the principle aspects of the present invention, well-known methods, procedures, components and circuitry in link server device 340 are not described in detail.

[0053] Link server device 340 includes a Land Control Protocol (LCP) interface 358 that couples to wired network 300, and a Wireless Control Protocol (WCP) interface 341 that couples to wireless network 320. A server module 310 is coupled between the LCP interface 358 and the WCP interface 341.

[0054] Server module 310 performs traditional server processing as well as protocol conversion processing from one communication protocol to another communication protocol. Message processor 316 is the component responsible for protocol conversions and associated tasks. In the case of protocol conversions (e.g. between HDTP and HTTP), the conversion is generally a data mapping process. It will be understood by those skilled in the art that WCP interface 341 can be replaced by other interface modules depending on the wireless networks and protocols used. The same is true of LCP interface 358 when the type of wired network and protocol vary.

[0055] Server module 310 also includes an account manager 312 and an account interface 314. Account manager 312 manages a plurality of user accounts, typically one for each of the mobile devices serviced by link server device 340. It should be understood that the user account information may be stored in another network server coupled to the link server device 340. In other words, the user accounts can be kept in a database that is physically placed in any computing device coupled to link server device 340 via a wired network.

[0056] Each of the mobile devices serviced by link server device 340 is assigned an identification (ID) or device ID. A device ID can be a phone number of the device or an IP address or a combination of an IP address and a port number, for example: 204.163.165.132:01905 where 204.163.165.132 is the IP address and 01905 is the port number. The device ID is further associated with a subscriber ID created and administrated by the carrier controlling link server device 340 as part of the procedures involved in activating a subscriber account for mobile device. The subscriber ID may be associated with, and utilized to, access the user specific files (e.g. 112 of Figure 1) associated with a particular user or device.

[0057] The subscriber ID may take the form of, for ex-

ample, 861234567-10900_pn.mobile.att.net by AT&T Wireless Service, and is a unique identification to a mobile device. The account manager 312 is responsible for creating a user account for a mobile device that allows for secure communications with link server device 340. In this case, account manager 312 ensures the proper level secure access for the serviced mobile device to services provided by link server device 340.

[0058] Link server device 340 also includes a processor 318 and storage resource 320 as the primary hardware components. Processor 318 performs operations under the control of the server module 310. It will be understood to those skilled in the art that link server device 340 may include one or more processors (e.g., processor 318), working memory (e.g., storage resource 320), buses, interfaces, and other components and that server module 310 represents one or more software modules loaded into the working memory of link server device 340 to perform designated functions. The same distinction is equally applicable to the client module and hardware components of the subject mobile devices.

[0059] Typically the landnet communication protocol (LCP) supported in landnet 300 may include Transmission Control Protocol (TCP), HyperText Transfer Protocol (HTTP) or Secure HyperText Transfer Protocol (HT-TPS), and the wireless communication protocol (WCP) may include (TCP), (HTTP) or (HTTPS), Handheld Device Transport Protocol (HDTP) or Wireless Session Protocol (WSP). In the case that LCP is different from WCP, server module 310 includes a mapping module (i.e. a mapper) responsible for mapping from one protocol to another so that a mobile device coupled to wireless network 320 can communicate with a device coupled to wired network 300.

[0060] Once the received speech signal is processed by voice recognition server system (not shown), a symbolic data file is generated and forwarded to link server device 340. The symbolic data file is received by message processor 316 via LCP interface 358. Message processor 316 converts the symbolic data file to a data format that may be optimally (in terms of the protocol requirements of the wireless network and the device characteristics of the requesting mobile device) transported on wireless network 320. The symbolic data file can be in a format comprehensible by message processor 316, for example, in a markup language (e.g. HTML) or a text file (e.g. ASCII), when received from the voice recognition server system. The processed symbolic data file, which may be reformatted so as to be more compatible with the requesting mobile device, is then forwarded to the requesting mobile device or to a designated third party device.

[0061] Referring to Figure 4, there are shown functional modules of an exemplary voice recognition server system 460 (which may correspond voice recognition server system 109 of Figure 1) that performs the following processes: 1) Speech Detection, 2) Speech Analy-

sis, 3) Pattern Matching and 4) Symbolic File Generation. During speech detection 462, Voice recognition server system 460 detects the presence of a speech signal at its input. Upon detection, the received speech signal goes through the speech analysis process 464, where it is reduced to quantifiable indices suitable for pattern matching. During the pattern matching stage 466, the quantifiable indices are compared to user voice templates (if using a template based voice recognition process) stored in storage device 480 that may include various language dictionaries and a plurality of user specific files. The symbolic data file is forwarded to link server device 340 via wired network 300 (see Figure 3) as previously described. It will be understood, by persons of ordinary skill in the art, that other voice recognition schemes (e.g. Fourier transforms or linear predictive coding (LPC)) may be used without deviating from the scope of the invention. Persons of ordinary skill in the art will also understand that the link server device (e.g. 106 of Figure 1) may perform the functions of the voice recognition server system (e.g. 109 of Figure 1).

[0062] Figure 5 illustrates a plurality of exemplary screen displays relating to the interaction of a mobile device requesting voice recognition services and a voice recognition server system. Initially screen display 500 allows a user to select between manual entry 504 and VR (voice recognition) assisted entry 508. User selections are indicated by selection indicator 512. In this example, VR assisted entry 512 may be selected by activating the softkey associated with softkey function identifier 516. This selection retrieves the contact information for the voice recognition server system providing service. In this example the contact information is comprised of a phone number (e.g. 650-555-7272). One of ordinary skill in the art will understand that the contact information may also be comprised of a Uniform Resource Identifier (URI) or similar unique identifier. Associated user and/or device identification information, utilized for accessing user specific files, may be transmitted in the background (e.g. using a separate data communication channel or the voice communication channel) or input by the user.

[0063] Upon retrieval of the voice recognition server system contact information 522, as shown in screen display 520, a voice channel may be established by activating the softkey associated with softkey function identifier 524 ("OK"). Screen display 530 illustrates types of information, which could be provided to the user of the requesting mobile device. Character string 532 provides the user with information relating to the status of establishing a communication session with the voice recognition server system providing service. Character string 534 provides the user with information relating to the settings utilized to process the user's request. This could be comprised of a simple character string (e.g. "Initializing Default Settings") or a plurality of interactive and non-interactive displays which allow a user to input selections (e.g. a language of choice). When the serv-

icing voice recognition server system is ready to receive input a prompt 536 ("begin speaking") is presented to the user. A user may end the input session by activating the softkey associated with softkey function identifier 538.

[0064] Voice recognition services may be configured to interact with particular applications resident on the requesting mobile device. For example, processed symbolic data files may be generated to serve as inputs for specific fields in an application such as an email. Additionally, once an active voice channel has been established for voice recognition services, the user may change the application using the service without having to secure and re-establish the voice communication channel. For example, the user may switch between an email program and a personal organizer. This feature reduces user cost and network congestion.

[0065] Referring now to Figure 6 and Figure 7, there are respectively illustrated process flowcharts that describe the operations of the mobile device and the voice recognition server according to one embodiment of the present invention. Both Figure 6 and Figure 7 should be understood in conjunction with Figure 1.

[0066] In accordance with the preferred embodiment of the present invention, a user desiring voice recognition services would initiate a request for the service by using the local user interface (e.g. by pressing a key). Generally the user would do so in conjunction with a desired task being performed using resident applications (e.g. email or web browsing). Information returned to the mobile device as a result of the request may be incorporated within a document associated with the task being performed.

[0067] The request process causes a voice channel to be established between the mobile device requesting service and the voice recognition server system providing the service. Once the voice channel is established and the user is queued to begin speaking, the user may begin an input interaction with the mobile device which may include physical input using the local user interface (e.g. a phone keypad) in addition to the voiced input. Upon completion of the initial input interaction with the mobile device the user may choose to maintain the open status of the voice channel open and perform another task or terminate the voice channel.

[0068] Figure 6 is a flow diagram, which illustrates the process 600 utilized by a mobile device (e.g. mobile devices 102 and 103) to interact with a remote voice recognition server system (e.g. voice recognition server system 109) from the perspective of the mobile device. At 604 a determination is made as to whether there is an active voice channel between the subject mobile device and the voice recognition server system providing services. This process usually occurs in the background under software control.

[0069] If there is an active voice channel, then the user is prompted to provide an input at 608 indicating whether the user desires the active voice channel to be

disabled. This would be the case where the user does not require VR services for the planned input interaction with the mobile device.

[0070] If the user decides to disable the voice channel then it is disabled at 612. The user then proceeds with physical input 628 using the device's user interface (e.g. the keypad). At 622 a decision is made as to whether user input (e.g. physical input 628) has been registered (e.g. input accepted by the device). If the user input is registered then it is processed at 632 and the user is prompted to provide an input at 636 indicating whether to continue the input session or terminate it. If the user selects termination then a determination is made as to the status of established voice channels/circuits at 640 (i.e. is the voice channel/circuit active). As was previously described, this check usually occurs in the background. In the sequence described above there is not an active voice channel so the process would be terminated.

[0071] If at 608 the user decides not to disable the voice channel, as would be the case where a user intends to utilize voice recognition services for a mobile device input interaction, then the user provides voice input 624 and physical input 628 and a determination is made at 622 as to whether the user's input has been registered. If the user's input has been registered then it is processed at 632 and the user is prompted to provide an indication at 636 as to whether the user desires to continue the input session or terminate it. If the user selects termination then a determination is made as to the status of any established voice channels/circuits at 640. Upon termination, active voice channels/circuits are secured. The process is then terminated.

[0072] If the user decides not to terminate the input session at 636 then the process returns to the beginning of process 600.

[0073] If a determination is made at 604 that there is not an active voice channel, then the user is prompted to provide an indication at 614 as to whether the user desires an active voice channel to be established. This would be the case where the user requires voice recognition services for an input interaction with the mobile device.

[0074] If at 614 the user requests a voice channel for the input interaction then one is established at 618. The user then provides voice input 624 and physical input 628 and a determination is made at 622 as to whether the user's input has been registered. If the user's input has been registered then it is processed at 632 and the user is prompted to provide an indication at 636 whether to continue the input session or terminate it. If the user selects termination then a determination is made, as previously described as to the status of any established voice channels/circuits at 640. Upon termination, active voice channels/circuits are secured. The process is then terminated.

[0075] If at 614 the user does not request a voice channel for the impending input interaction, as would be

the case where the user does not require voice recognition services, the user then proceeds with physical input 628 using the mobile device user interface (e.g. the keypad). At 622 a determination is made as to whether user input (e.g. physical input 628) has been registered. If the user input has been registered then it is processed at 632 and a decision is made at 636 whether to continue the input session or terminate it. The process is then terminated.

[0076] If the user decides not to terminate the input session at 636 then the process returns to the beginning of process 600.

[0077] If in any of these exemplary interactions described above, the user input is not registered at 622, then the user is prompted to provide an indication at 636 as to whether or not they desire to terminate the session with the voice recognition server system.

[0078] Once a voice channel between the voice recognition system providing service and a mobile device requesting service is established, the voice recognition server system may retrieve any user specific files associated with the user of the mobile device (e.g. language preferences, template files etc.) and use these to process the incoming voice input. The voice recognition server system then detects and processes incoming voice signals associated with the request for service. The incoming voice signal is converted into a symbolic data file using a template matching process, Fourier transform method, linear predictive coding scheme or any suitable voice recognition coding scheme and forwarded to the requesting mobile device (or a designated third party device) using a data communication channel that may include an intermediate server device (e.g. link server device 106 of Figure 1).

[0079] The symbolic data file may be in a format that is suitable for processing by the requesting mobile device (e.g. cHTML, WML or HDML) or may be in a format suitable for processing by an intermediate server device (e.g. HTML, WML, XML, ASCII etc.). In the latter case the intermediate server device may perform any conversion process required if any.

[0080] According to the principles of the present invention, a user interacting with a mobile device would be able to access remotely available voice recognition services based in a server device running a voice recognition application (e.g. a voice recognition server system). Software stored on the phone (e.g. a microbrowser) assists the user in this interaction by retrieving and managing contact information for the server device and by providing prompts and performing functions related to interactions with the voice recognition server system. Using this system and method, mobile devices having limited processing and storage capability have access to full featured voice recognition applications running on powerful computer workstations.

[0081] Figure 7 is a flow diagram, which illustrates the process 700 utilized by a voice recognition server system (e.g. voice recognition server system 109) to inter-

act with a mobile device (e.g. mobile device 102) from the perspective of the voice recognition server system. At 704 a determination is made (i.e. by a software process) as to whether a voice circuit/channel has been established between the voice recognition server system (e.g. voice recognition server system 109) and a mobile device requesting services (e.g. mobile device 102).

[0082] If at 704 it is determined that a voice circuit/channel has been established with a mobile device requesting services then another determination is made at 708 as to whether a speech signal has been detected. If a speech signal is detected at 708 the received speech input 716 is utilized to generate a symbolic data file at 712.

[0083] As previously stated, the symbolic data file is a file containing a plurality of letters, phonemes, words, figures, objects, functions, control characters or other conventional marks designating an object, quantity, operation, function, phoneme, word, phrase or any combination thereof having some relationship to the received speech signal as interpreted by the speech recognition system. Voice recognition systems generally use voice templates, Fourier Transform coding, or a linear predictive coding scheme to map the voiced input components to pre-stored symbolic building blocks. Examples of symbolic data files include ASCII files and binary data files.

[0084] The symbolic data file is then forwarded to the requesting mobile device (or designated third party device) at 720. At 724 it is determined whether a termination command has been received from the mobile device requesting services. If a termination command is received then the process is ended. If a termination command is not received then the process continues to look for an incoming speech signal at 708. If at 708 a speech signal is not received within a pre-determined time period then a determination is made at 728 as to whether a termination command has been received. If a termination command has been received then the process is terminated. Of course the system could have pre-determined time-outs or cycle limits that could result in process termination even if a termination command has not been received.

[0085] If at 704 it is determined that a voice circuit/channel has not been established with a mobile device requesting services then the voice recognition server system awaits the establishment of an active voice channel with a mobile device desiring voice recognition services.

[0086] According to the principles of the present invention, the voice recognition server system functions as an extension of the user interface of the mobile device. For example a user can choose to use voice recognition services for lengthy interactions that would under normal circumstances require considerable time and effort to input using the local user interface. In addition, since the resources of the mobile device do not limit the voice recognition application used, the user can

be provided access to a vast vocabulary.

[0087] The advantages of the present invention are numerous. Different implementations may yield one or more of the following advantages. One advantage of the present invention is that users of certain mobile devices (e.g. devices with limited processing and storage capability) are enabled to utilize a fully functional voice recognition application running on a remote server device to augment the standard device user interface.

[0088] Another advantage of the present invention is that since the voice recognition application utilized is not restricted by the processing and storage limitations of the mobile device, the user may be provided with the functionality of a full feature voice recognition application running on a more powerful computer. Advantages associated with this functionality include providing the user with multiple language dictionaries having large vocabularies and personalized dictionaries. Additionally, since the voice recognition application is not stored on the mobile device, there is little or no impact on the per-unit cost of the mobile device. Still another advantage of the present invention is that the carriers providing this service can charge the users a small service fee for access to it.

[0089] Yet another advantage of the present invention is that a user can utilize voice recognition services and the local user interface (e.g. a phone keypad) concurrently, thus providing the user with maximum flexibility. For example, the user can input a voice signal and intermingle symbols from the local user interface.

[0090] The many features and advantages of the present invention are apparent from the written description, and thus, it is intended by the appended claims to cover all such features and advantages of the invention.

Further, since numerous modifications and changes will readily occur to those skilled in the art, it is not desired to limit the invention to the exact construction and operation as illustrated and described. Hence, all suitable modifications and equivalents may be resorted to as falling within the scope of the invention.

Claims

- 45 1. A method of providing voice recognition services to a wireless communication device having a display screen and a user interface, the method comprising:
50 receiving a request from the wireless communication device for voice recognition services at a server device running a voice recognition application;
55 retrieving a voice input signal associated with the request from a first communication path; converting the voice input signal into a symbolic data file using the voice recognition application; and

forwarding the symbolic data file to the wireless communication device using a second communication path.

2. A method as recited in claim 1, wherein the first communication path is established on a wireless communication network. 5

3. A method as recited in claim 1 or 2, wherein the symbolic data file is a markup language file. 10

4. A method as recited in any preceding claim, wherein the second communication path includes a link server device connected to the server device running the voice recognition application by a wired network using a first communication protocol and to the wireless communication device by a wireless network using a second communication protocol. 15

5. A method of providing voice recognition services to a wireless communication device having a display screen and a user interface, the method comprising:-

retrieving contact information for a server device running a voice recognition application; generating a request for voice recognition services from the server device associated with the retrieved contact information; forwarding the request for voice recognition services to the server device associated with the retrieved contact information; establishing a voice communication channel between the wireless communication device and the server device associated with the retrieved contact information; receiving input from a user using the wireless communication device, at least a portion of the input including a voice component; and transmitting the user input to the server device for processing by the voice recognition application. 25

6. A method as recited in claim 5, further comprising:-

receiving a symbolic data file from the server device associated with the retrieved contact information, the symbolic data file including the processed output of voice recognition processing of the user input by the server device; processing the received symbolic data file using the local resources of the wireless communication device; and displaying at least a portion of the processed symbolic data file to the user for review and modification. 30

7. A method as recited in claim 6, wherein the symbol- 35

ic data file is an ASCII formatted data file.

8. A computer readable medium on which is encoded computer program code for generating a request for voice recognition services for a wireless communication device, the medium comprising:-

computer program code for retrieving contact information for a server device providing voice recognition services; computer program code for generating a request for voice recognition services from the server device associated with the retrieved contact information; computer program code for receiving voice input from a user of the wireless communication device, the input being associated with the request for voice recognition services; and computer program code for establishing a voice communication session between the wireless communication device and the server device for the purpose of transmitting a voice signal to the server device for voice recognition processing. 40

9. A computer readable medium on which is encoded computer program code for providing voice recognition services to a wireless communication device, the medium comprising:-

computer program code for processing a request for voice recognition services received from the mobile device; computer program code for receiving a voice input associated with the request for voice recognition services; computer program code for converting the received voice input into a symbolic data file; and computer program code for forwarding the symbolic data file to the mobile device originating the request. 45

10. A wireless communication system providing voice recognition services, the system comprising:-

a wireless communication device providing voice input for voice recognition processing on a first communication path and receiving a symbolic data file representing the processed voice input on a second communication path; and a server device running a voice recognition application receiving voice input from the wireless communication device on the first communication path, converting the received voice input into a symbolic data file and forwarding the symbolic data file to the wireless device using the second communication path. 50

55

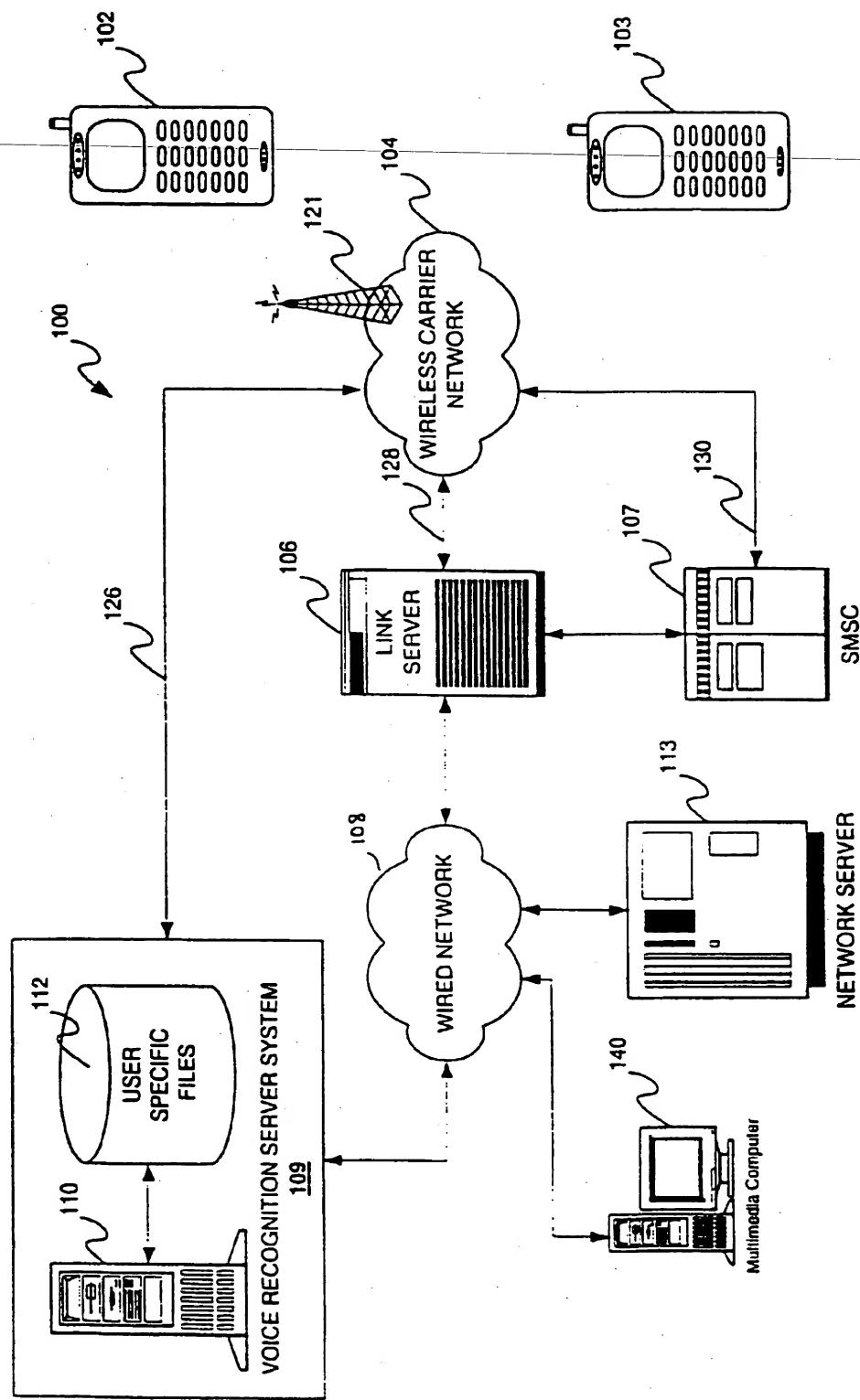


Figure 1

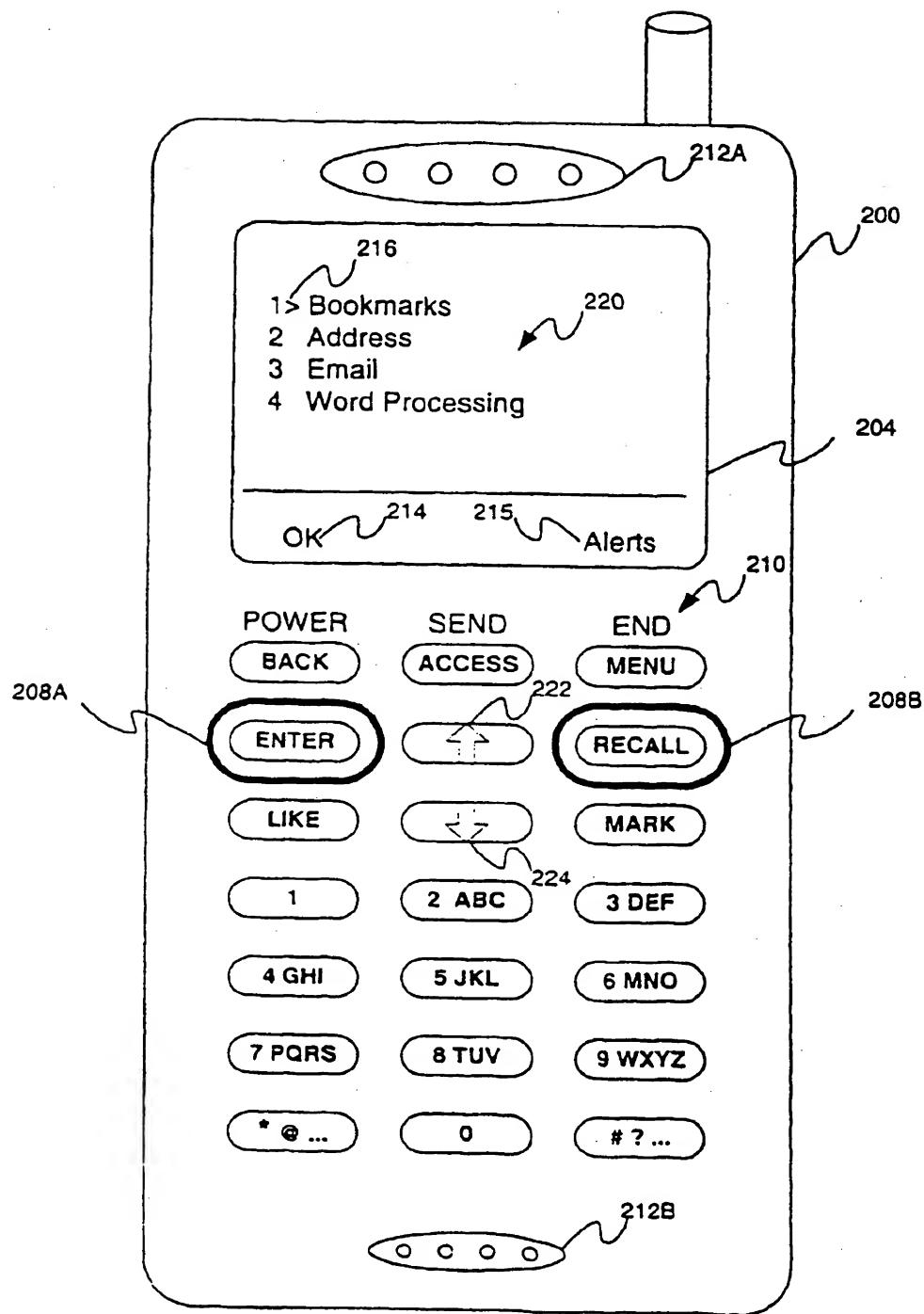


Figure 2A

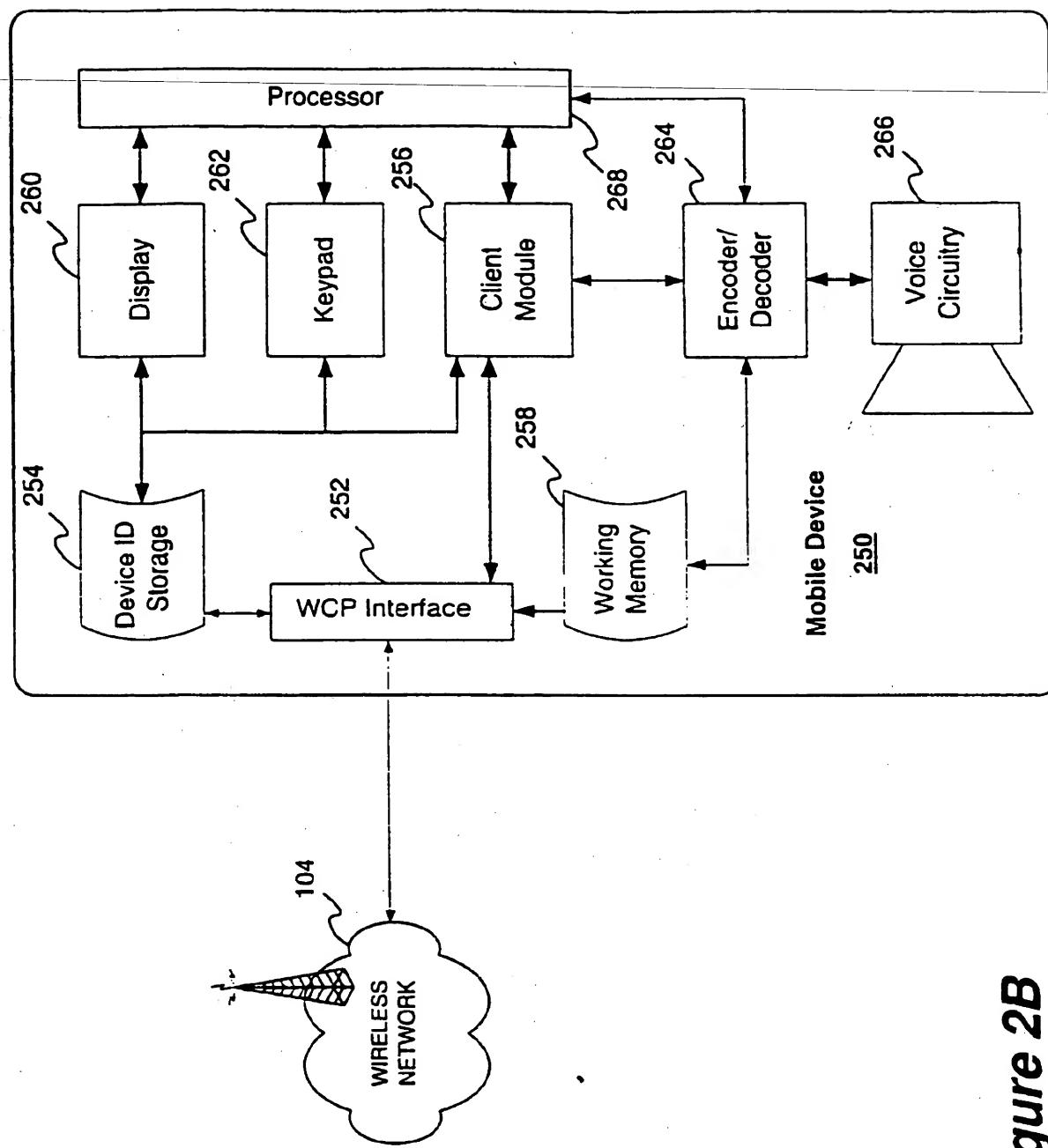


Figure 2B

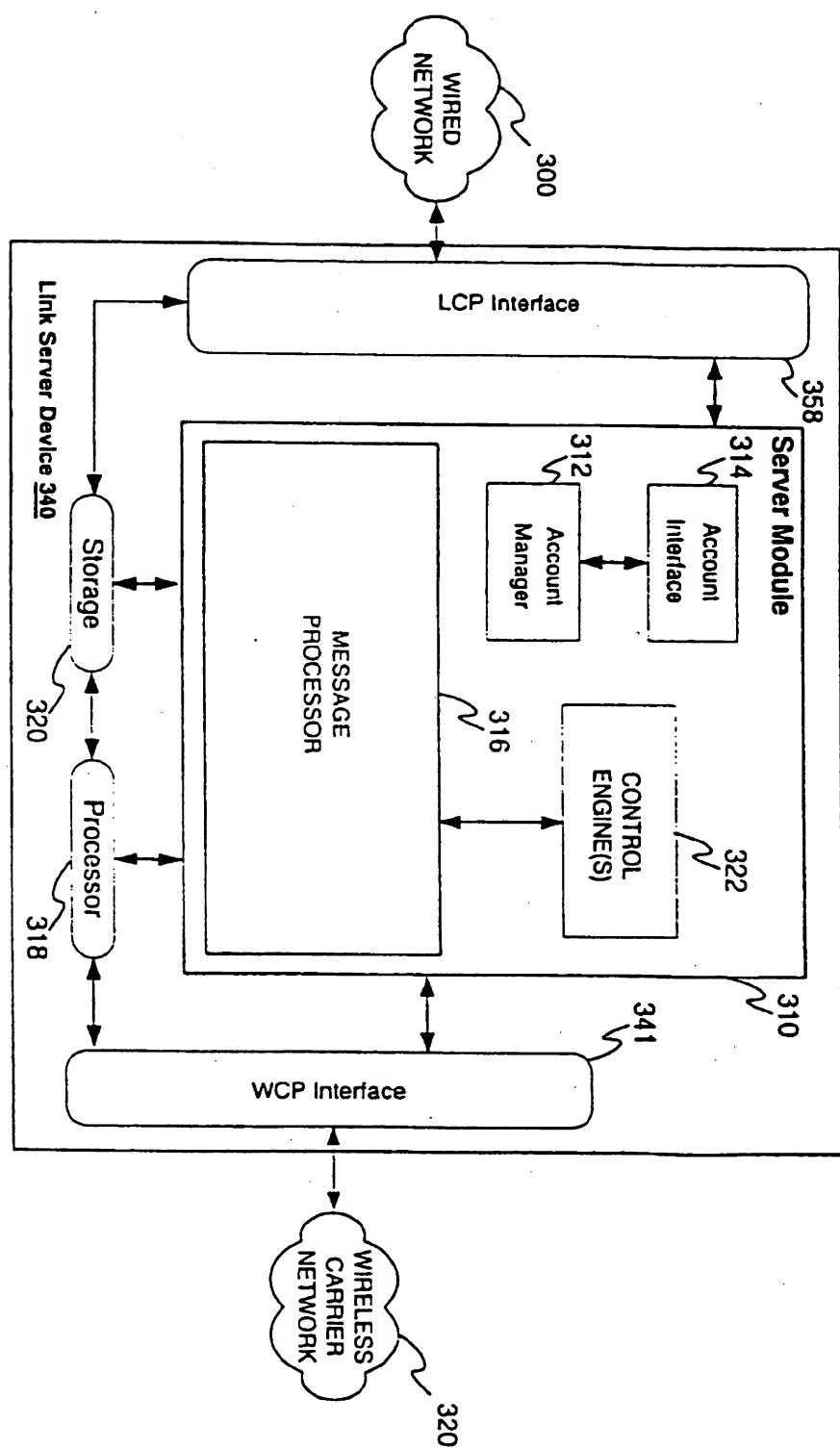


Figure 3

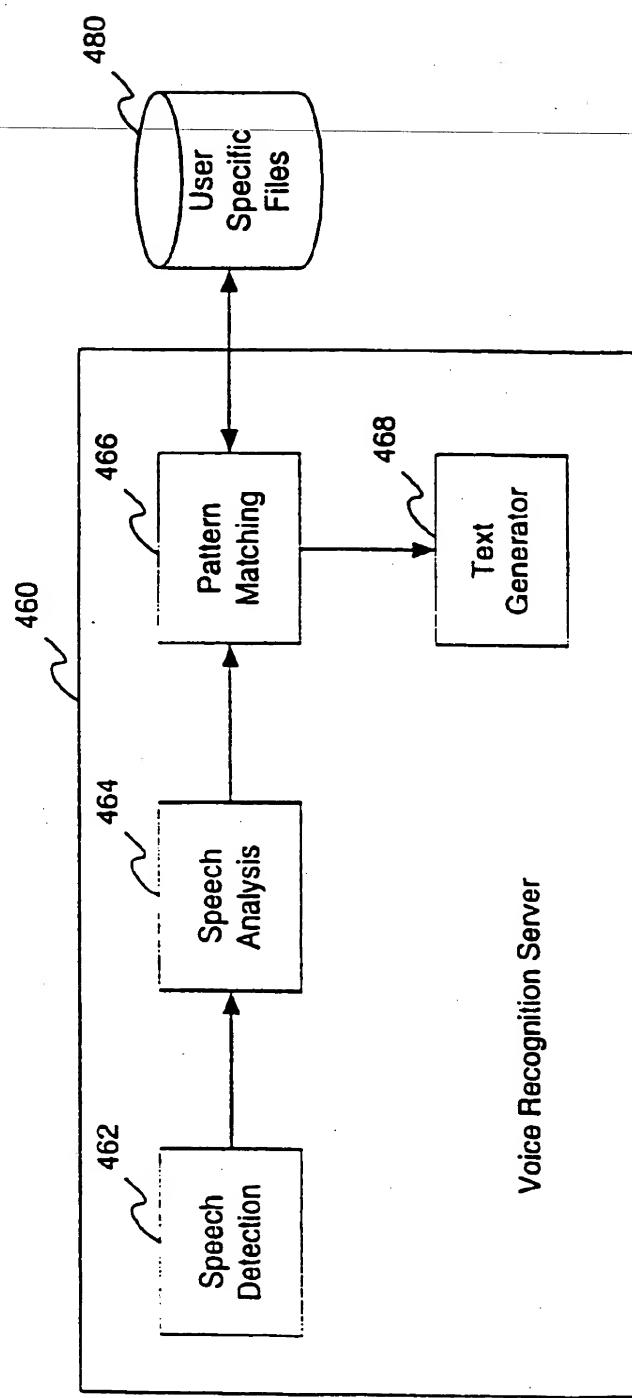


Figure 4

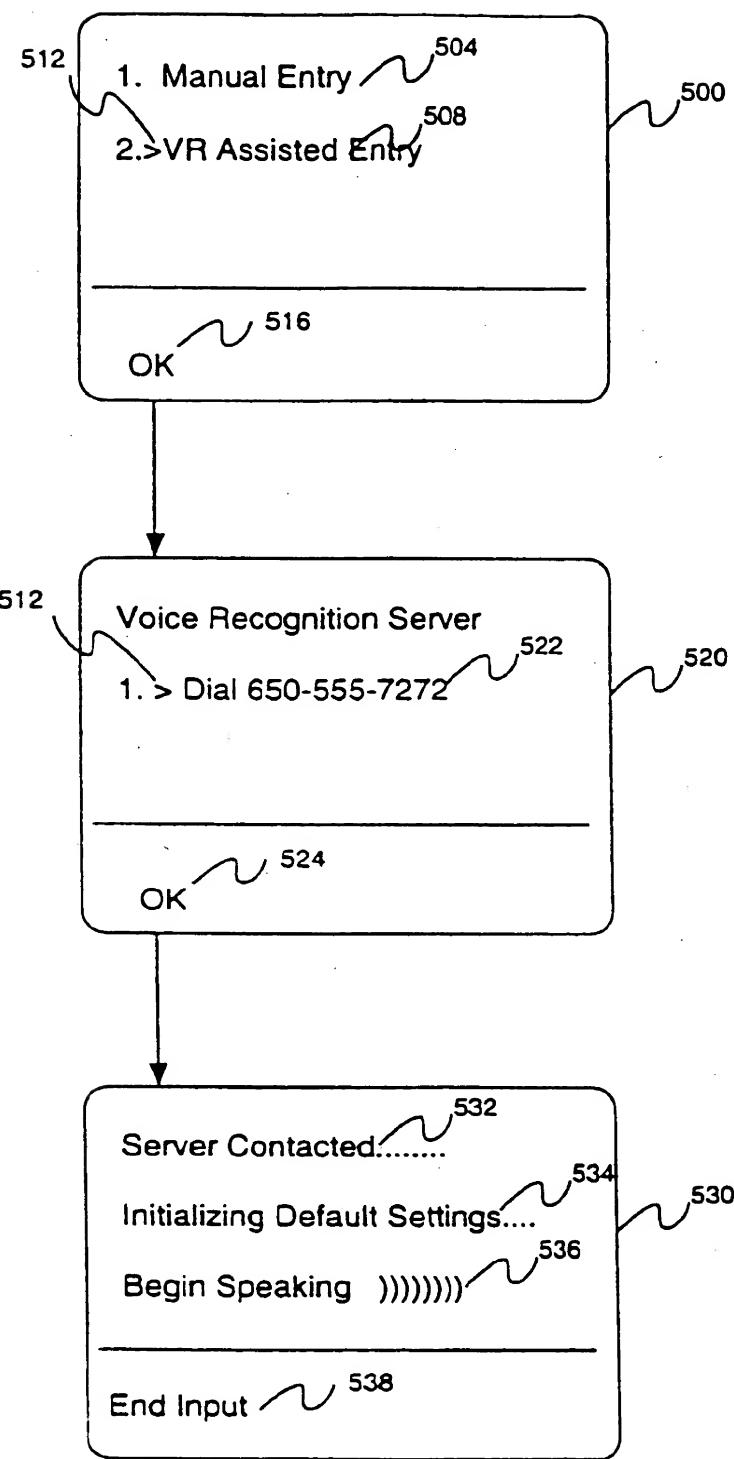
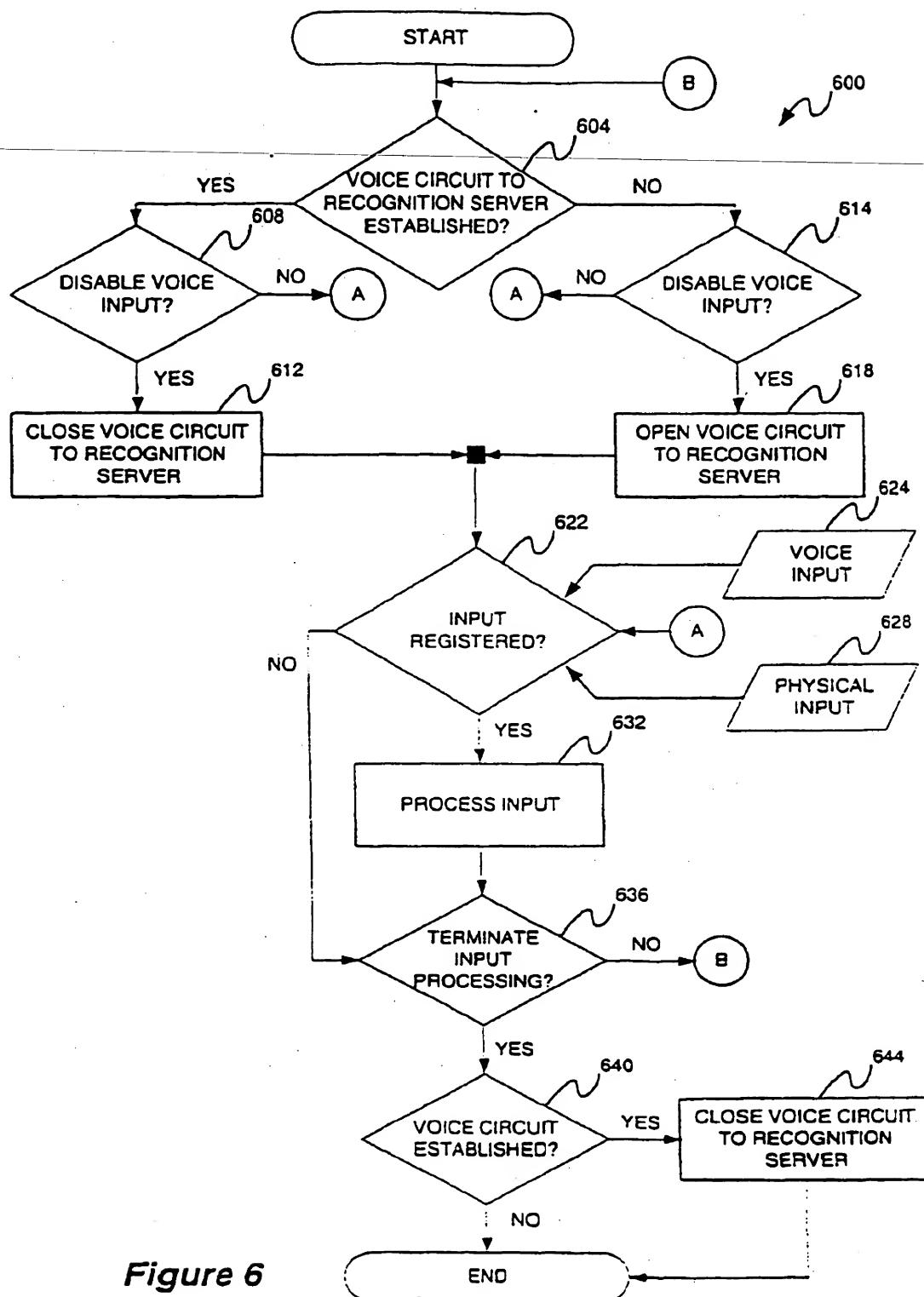


Figure 5



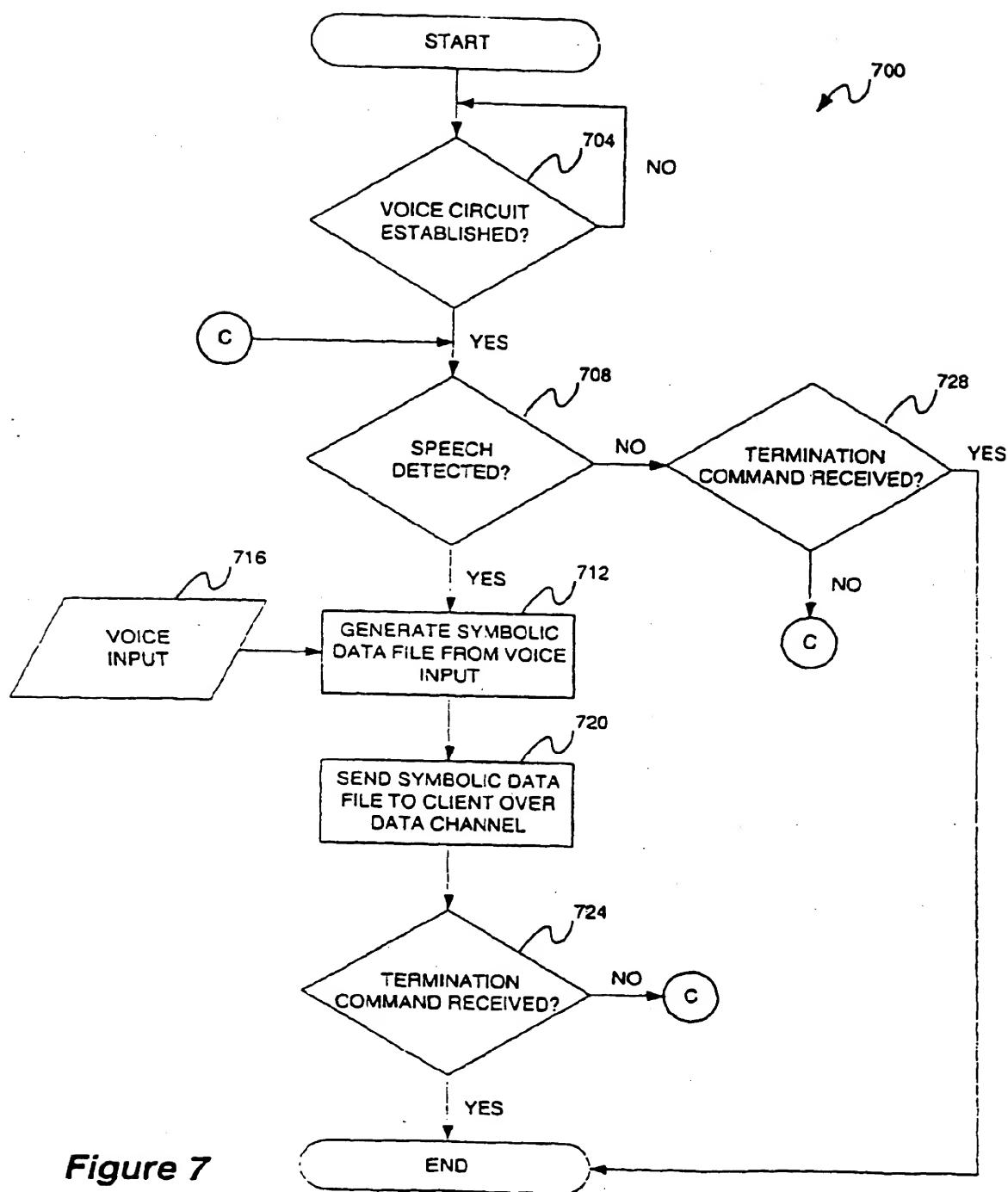


Figure 7

This page blank (up to)